



ENTERED

PCT

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/031,167

DATE: 12/03/2002

TIME: 13:25:46

Input Set : A:\seq INSERM-Final-Aug-02.txt

Output Set: N:\CRF4\12032002\J031167.raw

3 <110> APPLICANT: WAHBI, Kamal et al.
 5 <120> TITLE OF INVENTION: Nucleic acids coding for peptides having the biological activity of sorbin

7 <130> FILE REFERENCE: P07500US00/BAS

9 <140> CURRENT APPLICATION NUMBER: 10/031,167

C--> 10 <141> CURRENT FILING DATE: 2002-08-29

12 <150> PRIOR APPLICATION NUMBER: PCT/FR00/02076

13 <151> PRIOR FILING DATE: 2000-07-19

15 <150> PRIOR APPLICATION NUMBER: FR 99/09406

16 <151> PRIOR FILING DATE: 1999-07-20

18 <160> NUMBER OF SEQ ID NOS: 20

20 <170> SOFTWARE: PatentIn Ver. 2.1

22 <210> SEQ ID NO: 1

23 <211> LENGTH: 474

24 <212> TYPE: DNA

25 <213> ORGANISM: swine

27 <400> SEQUENCE: 1

```

28 atgagagcag caacaccttt gcagacagtt gaccggccga aggactggta caagaccatg 60
29 ttttaagcaaa tccacatggt gcacaagcca gatgatgaca cagacatgta taatactcct 120
30 tatacatata atgcaggcct gtacaactca ccctacagtg ctcagtcaca tcctgctgcc 180
31 aagaccccaga cctacagacc cctctccaaa agccactctg acaatggcac cgacgccttt 240
32 aaggatgctt cctcacctgt ccctcccca catgttcctc ctccagtccc acctctgcga 300
33 ccaagagatc ggtcttcaac agaaaagcat gactgggatc ctccagacag aaaagtggac 360
34 acgagaaaaat ttcgatcgga gccacggtct atttttgaat acgagcctgg gaagtcaccc 420
35 atcctgcagc acgaacgacc cgtcacgaaa ccgcaagcag ggcgcgctaa ggtc 474

```

38 <210> SEQ ID NO: 2

39 <211> LENGTH: 153

40 <212> TYPE: PRT

41 <213> ORGANISM: pig

43 <220> FEATURE:

44 <221> NAME/KEY: MOD_RES

45 <222> LOCATION: (153)

46 <223> OTHER INFORMATION: AMIDATION

48 <400> SEQUENCE: 2

```

49 Met Arg Ala Ala Thr Pro Leu Gln Thr Val Asp Arg Pro Lys Asp Trp
50   1           5           10           15
52 Tyr Lys Thr Met Phe Lys Gln Ile His Met Val His Lys Pro Asp Asp
53           20           25           30
55 Asp Thr Asp Met Tyr Asn Thr Pro Tyr Thr Tyr Asn Ala Gly Leu Tyr
56           35           40           45
58 Asn Ser Pro Tyr Ser Ala Gln Ser His Pro Ala Ala Lys Thr Gln Thr
59           50           55           60
61 Tyr Arg Pro Leu Ser Lys Ser His Ser Asp Asn Gly Thr Asp Ala Phe
62 65           70           75           80

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/031,167

DATE: 12/03/2002
TIME: 13:25:46

Input Set : A:\seq INSERM-Final-Aug-02.txt
Output Set: N:\CRF4\12032002\J031167.raw

```

64 Lys Asp Ala Ser Ser Pro Val Pro Pro Pro His Val Pro Pro Pro Val
65                      85                      90                      95
67 Pro Pro Leu Arg Pro Arg Asp Arg Ser Ser Thr Glu Lys His Asp Trp
68                      100                      105                      110
70 Asp Pro Pro Asp Arg Lys Val Asp Thr Arg Lys Phe Arg Ser Glu Pro
71                      115                      120                      125
73 Arg Ser Ile Phe Glu Tyr Glu Pro Gly Lys Ser Ser Ile Leu Gln His
74                      130                      135                      140
76 Glu Arg Pro Val Thr Lys Pro Gln Ala
77 145                      150
81 <210> SEQ ID NO: 3
82 <211> LENGTH: 492
83 <212> TYPE: DNA
84 <213> ORGANISM: Homo sapiens
86 <400> SEQUENCE: 3
87 atgaaagcaa caacaccttt gcagacagtc gaccggccca aggactggta caagacgatg 60
88 tttaagcaaa ttcacatggt gcacaagccg gatgatgaca cagacatgta taatactcct 120
89 acacctcaca tgaaatatac atacaatgca ggtctgtaca acccacccta cagtgtcag 180
90 tcacaccctg ctgcaaagac ccaaacctac agacctcttt ccaaagcca ctccgacaac 240
91 agccccaatg cctttaagga tgcgtcctcc ccagtgcctc ccccatgt tccacctcca 300
92 gtcccgccgc ttcgaccaag agatcgggtc tcaacagaaa agcatgactg ggatcctcca 360
93 gacagaaaag tggacacaag aaatttcggg tctgagccaa ggagtatttt tgaatacgag 420
94 cctgggaagt catccatcct gcagcacgaa cgaccgtca cgaaaccgca agcagggcgc 480
95 cgtgataagt cc 492
98 <210> SEQ ID NO: 4
99 <211> LENGTH: 158
100 <212> TYPE: PRT
101 <213> ORGANISM: Homo sapiens
103 <220> FEATURE:
104 <221> NAME/KEY: MOD_RES
105 <222> LOCATION: (158)
106 <223> OTHER INFORMATION: AMIDATION
108 <400> SEQUENCE: 4
109 Met Lys Ala Thr Thr Pro Leu Gln Thr Val Asp Arg Pro Lys Asp Trp
110 1 5 10 15
112 Tyr Lys Thr Met Phe Lys Gln Ile His Met Val His Lys Pro Asp Asp
113 20 25 30
115 Asp Thr Asp Met Tyr Asn Thr Pro Thr Pro His Met Lys Tyr Thr Tyr
116 35 40 45
118 Asn Ala Gly Leu Tyr Asn Pro Pro Tyr Ser Ala Gln Ser His Pro Ala
119 50 55 60
121 Ala Lys Thr Gln Thr Tyr Arg Pro Leu Ser Lys Ser His Ser Asp Asn
122 65 70 75 80
124 Ser Pro Asn Ala Phe Lys Asp Ala Ser Ser Pro Val Pro Pro Pro His
125 85 90 95
127 Val Pro Pro Pro Val Pro Pro Leu Arg Pro Arg Asp Arg Ser Ser Thr
128 100 105 110
130 Glu Lys His Asp Trp Asp Pro Pro Asp Arg Lys Val Asp Thr Arg Asn
131 115 120 125

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/031,167

DATE: 12/03/2002

TIME: 13:25:46

Input Set : A:\seq INSERM-Final-Aug-02.txt

Output Set: N:\CRF4\12032002\J031167.raw

```

133 Phe Gly Ser Glu Pro Arg Ser Ile Phe Glu Tyr Glu Pro Gly Lys Ser
134      130                      135                      140
136 Ser Ile Leu Gln His Glu Arg Pro Val Thr Lys Pro Gln Ala
137 145                      150                      155
141 <210> SEQ ID NO: 5
142 <211> LENGTH: 1794
143 <212> TYPE: DNA
144 <213> ORGANISM: Homo sapiens
146 <400> SEQUENCE: 5
147 atgaaagcaa caacaccttt gcagacagtc gaccggccca aggactggta caagacgatg 60
148 tttaagcaaa ttcacatggt gcacaagccg gatgatgaca cagacatgta taatactcct 120
149 acacctcaca tgaaatatac atacaatgca ggtctgtaca acccacccta cagtgtctcag 180
150 tcacaccctg ctgcaaagac ccaaacctac agacctcttt ccaaaagcca ctccgacaac 240
151 agccccaatg cctttaagga tgcgtcctcc ccagtgcctc cccacatgt tccacctcca 300
152 gtcccgccgc ttcgaccaag agatcgggtc tcaacagaaa agcatgactg ggatcctcca 360
153 gacagaaaag tggacacaag aaatttcggg tctgagccaa ggagtatttt tgaatacgag 420
154 cctgggaagt catccatcct gcagcacgaa cgaccctct accagtcttc catagacaga 480
155 agcttggaag gaccagcag ctctgcaagc atggcggtg actttagaaa acggaggaag 540
156 agtgaacctg cagtgggccc gccaggggc ttgggggatc acagttcaag caggaccagc 600
157 cccggccggg cagacctccc aggatcaagt tccaccttta ccacgtcttt cattagttct 660
158 tctccttctc ctccctcgag agcacaaggt ggggatgata gcaaaatgtg tccgcccctt 720
159 tgcagttact cggggctcaa tggctcggcc tctagttagt tagagtgtg cggcgttat 780
160 agaaggcact tggacgtccc ccaggactct caaagggccca tcactttcaa gaacggctgg 840
161 caaatggccc ggcaaatgca agagatctgg agtagcactg aagaggcggg tcccccaaa 900
162 atcaaatcac gaagctgtga cgatctcctg aatgatgact gcggcagctt ccagaccct 960
163 aaaaccaagt cagaaagcat gggttctctg ttatgtgacg aaggctccaa agagagcgac 1020
164 cccatgacgt ggacttcccc ctacatcccg gaagtgtgca ggaacagcag agaattcatg 1080
165 tttaagcaaa tggatattcg tggatctctt ggatggagga ccattttgga aagtgtctaa 1140
166 ggaatatcta taatgagtga ggaatctatg agaaagatgt aaagtgtgaa acgtaaattt 1200
167 tttggtttag tagatgatca ctgatttaaa tgtataacag agtagatgcc cccccctca 1260
168 aaaacgcata accccccctt taccctgaca tttagctttg aatatgcaca aaatagtttg 1320
169 tgggtagaat agaaccctat gtctgaaagt atatgtgttg ggatttcac ccatatatgg 1380
170 tggtagccgc caactcagag ataggctcgt ctggttagatt ctcaacaaca aaatgtataa 1440
171 cacaagcttg aattcatgtt taagcaaata aaaataatgt gggagactgg acagaggtca 1500
172 gggacccag ggtgccaaagt gtagctcaga gtcaccattg gtgaatcgct tcatctccat 1560
173 gtggaactaa atgcaactaa gtgatttctt aggtttccc cagtcattct tagtgaaaat 1620
174 atggacttcc cacatcaatt ctgagtcact ttcttccac ctggaatgat taccattttt 1680
175 ctcatagtca gtgtatgcag cagcatatac cctcatttgc ctttgggtac attcctgagt 1740
176 caaaatgtat aacacaaggt cacgaaaccg caagcagggc gccgtgataa gtcc 1794
179 <210> SEQ ID NO: 6
180 <211> LENGTH: 21
181 <212> TYPE: DNA
182 <213> ORGANISM: Homo sapiens
184 <400> SEQUENCE: 6
185 ccggtcacga aaccgcaagc a
188 <210> SEQ ID NO: 7
189 <211> LENGTH: 30
190 <212> TYPE: DNA
191 <213> ORGANISM: Homo sapiens

```

21

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/031,167

DATE: 12/03/2002

TIME: 13:25:46

Input Set : A:\seq INSERM-Final-Aug-02.txt
 Output Set: N:\CRF4\12032002\J031167.raw

```

193 <400> SEQUENCE: 7
194 cacgaacgac ccgtcacgaa accgcaagca 30
197 <210> SEQ ID NO: 8
198 <211> LENGTH: 120
199 <212> TYPE: DNA
200 <213> ORGANISM: Homo sapiens
202 <400> SEQUENCE: 8
203 cctccagaca gaaaagtgga cacaagaaat ttcgggtctg agccaaggag tatttttgaa 60
204 tacgagcctg ggaagtcac ccatcctgcag cacgaacgac ccgtcacgaa accgcaagca 120
207 <210> SEQ ID NO: 9
208 <211> LENGTH: 7
209 <212> TYPE: PRT
210 <213> ORGANISM: Homo sapiens
212 <220> FEATURE:
213 <221> NAME/KEY: MOD_RES
214 <222> LOCATION: (7)
215 <223> OTHER INFORMATION: AMIDATION
217 <400> SEQUENCE: 9
218 Pro Val Thr Lys Pro Gln Ala
219 1 5
223 <210> SEQ ID NO: 10
224 <211> LENGTH: 10
225 <212> TYPE: PRT
226 <213> ORGANISM: Homo sapiens
228 <220> FEATURE:
229 <221> NAME/KEY: MOD_RES
230 <222> LOCATION: (10)
231 <223> OTHER INFORMATION: AMIDATION
233 <400> SEQUENCE: 10
234 His Glu Arg Pro Val Thr Lys Pro Gln Ala
235 1 5 10
239 <210> SEQ ID NO: 11
240 <211> LENGTH: 40
241 <212> TYPE: PRT
242 <213> ORGANISM: Homo sapiens
244 <220> FEATURE:
245 <221> NAME/KEY: MOD_RES
246 <222> LOCATION: (40)
247 <223> OTHER INFORMATION: AMIDATION
249 <400> SEQUENCE: 11
250 Pro Pro Asp Arg Lys Val Asp Thr Arg Asn Phe Gly Ser Glu Pro Arg
251 1 5 10 15
253 Ser Ile Phe Glu Tyr Glu Pro Gly Lys Ser Ser Ile Leu Gln His Glu
254 20 25 30
256 Arg Pro Val Thr Lys Pro Gln Ala
257 35 40
261 <210> SEQ ID NO: 12
262 <211> LENGTH: 17
263 <212> TYPE: DNA

```

RAW SEQUENCE LISTING

DATE: 12/03/2002

PATENT APPLICATION: US/10/031,167

TIME: 13:25:46

Input Set : A:\seq INSERM-Final-Aug-02.txt

Output Set: N:\CRF4\12032002\J031167.raw

```

264 <213> ORGANISM: Artificial sequence
266 <220> FEATURE:
267 <223> OTHER INFORMATION: Description of the artificial sequence: primers
268     used for the RT-PCR
270 <220> FEATURE:
271 <221> NAME/KEY: misc_feature
272 <222> LOCATION: (9)..(9)
273 <223> OTHER INFORMATION: n=(a or c or t or g)
276 <400> SEQUENCE: 12
W--> 277 aargayacnt ayaarac                                17
280 <210> SEQ ID NO: 13
281 <211> LENGTH: 17
282 <212> TYPE: DNA
283 <213> ORGANISM: artificial sequence
285 <220> FEATURE:
286 <223> OTHER INFORMATION: Description of the artificial sequence: primers
287     used for the RT-PCR
290 <400> SEQUENCE: 13
291 cggccgaagg actggta                                17
294 <210> SEQ ID NO: 14
295 <211> LENGTH: 18
296 <212> TYPE: DNA
297 <213> ORGANISM: artificial sequence
299 <220> FEATURE:
300 <223> OTHER INFORMATION: Description of the artificial sequence: primers
301     used for the RT-PCR
303 <400> SEQUENCE: 14
304 acaagccgag atgatgac                                18
307 <210> SEQ ID NO: 15
308 <211> LENGTH: 22
309 <212> TYPE: DNA
310 <213> ORGANISM: artificial sequence
312 <220> FEATURE:
313 <223> OTHER INFORMATION: Description of the artificial sequence: primers
314     used for the RT-PCR
316 <400> SEQUENCE: 15
317 gtcttcaaca gaaaagcatg ac                            22
320 <210> SEQ ID NO: 16
321 <211> LENGTH: 17
322 <212> TYPE: DNA
323 <213> ORGANISM: artificial sequence
325 <220> FEATURE:
326 <223> OTHER INFORMATION: Description of the artificial sequence: primers
327     used for the RT-PCR
329 <220> FEATURE:
330 <221> NAME/KEY: misc_feature
331 <222> LOCATION: (3)..(3)
332 <223> OTHER INFORMATION: n=(a or c or t or g)
335 <400> SEQUENCE: 16

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/031,167

DATE: 12/03/2002
TIME: 13:25:47

Input Set : A:\seq INSERM-Final-Aug-02.txt
Output Set: N:\CRF4\12032002\J031167.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:12; N Pos. 9
Seq#:16; N Pos. 3

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 5

VERIFICATION SUMMARY

DATE: 12/03/2002

PATENT APPLICATION: US/10/031,167

TIME: 13:25:47

Input Set : A:\seq INSERM-Final-Aug-02.txt

Output Set: N:\CRF4\12032002\J031167.raw

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:277 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0
L:336 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:0